

Supplemental material: additional simulation studies

Analyzing partially paired data: when can the unpaired portion(s) be safely ignored?
 Qianya Qi, Li Yan, and Lili Tian

The Appendix in the paper demonstrates that the asymptotic power is equivalent to the theoretical power when variances are known under normality, under certain regularity conditions. Additionally, we perform further simulation studies to demonstrate that the above findings are indeed valid at finite sample sizes. We also explore scenarios when normality assumption is not met.

Table S1 shows the estimated power of paired t-test T_p , unweighted and weighted versions of combination test T_1 given $\gamma_1 = n_2/n_1$, $\nu^2 = \sigma_X^2/\sigma_Y^2$, and ρ for bivariate normal data when variances are unknown. P-values in T_1 come from paired t-test T_p and two-sample t-test T_{up} . Sample size ratio γ_1 is set as $\frac{1}{4}, \frac{1}{2}, 1, 2$ where $n_1 = 32$, variance ratio ν^2 is set as $\frac{1}{3}, 1, 3$ where $\sigma_Y^2 = 2$, and correlation between paired samples ρ is set as $0.1, 0.2, \dots, 0.9$, the true mean difference is set as 0.4. Each scenario is repeated 5000 times. Theoretical maximum ρ 's for the unweighted or weighted P-value pooling test T_1 being more powerful than the naive paired test T_p when variances are known are shown in the ρ_T column. The estimated maximum ρ 's that satisfy combination tests being more powerful than naive paired tests are indicated by “*” or “-” or “+” in the table. For example, when $\gamma_1 = 1/4, \nu^2 = 1$, and $\rho = 0.1$, estimated power for naive paired test is 0.324, for unweighted combination test it is 0.313, the “-” in the table indicates that the estimated maximum correlation for the unweighted test T_1 being more powerful than the naive paired test T_p is less than 0.1. The estimated power for weighted combination test is 0.381 when ρ is 0.3, so the “*” indicates that the estimated maximum correlation is between 0.3 and 0.4. The estimated maximum correlations are very close to the theoretical maximum correlations which are 0.05 and 0.41 for unweighted and weighted combination tests respectively. In summary, from Table S1, the estimated maximum correlations for partially paired data with incompleteness in single arm agree with the values in Table 3 in the manuscript when variances are known.

Table S2 shows the estimated power of paired t-test T_p , unweighted and weighted versions of combination test T_2 given $\gamma_1 = n_2/n_1, \gamma_2 = n_3/n_2 = 1, \nu^2 = \sigma_X^2/\sigma_Y^2$, and ρ for bivariate normal data when variances are unknown for partially paired data with incompleteness in both arms. Parameter settings are the same as Table S1. Same conclusion can be made that the estimated maximum correlations for partially paired data with incompleteness in both arms agree with the theoretical values when variances are known.

Table S3 shows the estimated power of paired test T_p , unweighted and weighted versions of combination test T_1 given $\gamma_1 = n_2/n_1, \nu^2 = \sigma_X^2/\sigma_Y^2$, and ρ for bivariate logistic distribution for partially paired data with incompleteness in single arm. Sample sizes are set as (40, 10), (40, 20), (40, 40), (40, 80), corresponding to sample size ratio $\frac{1}{4}, \frac{1}{2}, 1, 2$, respectively. The marginal distributions are $Y \sim Logistic(0, 1)$, $X \sim Logistic(0.4, \sigma_X^2)$, where $\sigma_X^2 = \frac{1}{3}, 1, 3$ corresponding to variance ratio $\frac{1}{3}, 1, 3$, respectively. Correlation ρ is set as $0.1, 0.2, \dots, 0.9$, and each scenario is repeated 5000 times. As data is not normally distributed, and sample size here is only small or moderate, slight differences in estimated maximum correlations comparing to the theoretical

results are expected. From Table S3, the estimated maximum correlations in all scenarios generally match with the findings in the manuscript, with the differences no more than 0.2, except the scenario $\gamma_1 = 1/4$, $\nu^2 = 3$ for unweighted combination test: the theoretical maximum correlation is 0.38, but the estimated correlation is larger than 0.9. However, when $\gamma_1 = 1/4$ and $\nu^2 = 3$, the sample size in the unpaired part is only 10, estimated type I error from simulation indicates that the unweighted combination test has inflated type I error, thus may lead to a larger power.

A big advantage of P-value pooling method is that the P-values can come from any tests, so if data is not normally distributed, other tests such as non-parametric tests could be used. Simulations using P-values from Wilcoxon signed-rank test for the paired samples and Wilcoxon-Mann-Whitney test for the unpaired samples are also conducted, results are not shown here, but the findings are similar to Table S3.

Table S4 shows the estimated power of paired test T_p , unweighted and weighted versions of combination test T_2 given $\gamma_1 = n_2/n_1$, $\gamma_2 = n_3/n_2 = 1$, $\nu^2 = \sigma_X^2/\sigma_Y^2$, and ρ for bivariate logistic distribution for partially paired data with incompleteness in both arms. Parameter settings are the same as Table S3. Similarly to Table S3, the estimated maximum correlations generally agree with the theoretical values in the manuscript, with differences no more than 0.2, except the scenarios $\gamma_1 = 1/4$ when unpaired sample size is only 10. Similar trends are also confirmed (results are not shown here) when test statistics in P-value pooling tests and naive paired tests are from non-parametric tests.

In summary, the asymptotic power calculations in Appendix and simulation studies for data with incompleteness in either single arm or both arms, demonstrate that the results presented in the manuscript for normality with known variances may be used as crude guides for choices of tests when dealing with partially paired data.

Table S1: With incompleteness in single arm under bivariate normal distribution with unknown variances:
power of naive paired test T_p , unweighted and weighted combination tests T_1 given $\gamma_1 = n_2/n_1$,
 $\nu^2 = \sigma_X^2/\sigma_Y^2$, and ρ .

γ_1	ν^2	ρ_T	Method	ρ								
				0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1/4	1/3	-	P	0.410	0.432	0.476	0.503	0.569	0.616	0.691	0.800	0.890
		0	UW	0.394-	0.407	0.434	0.447	0.489	0.506	0.562	0.637	0.702
		0.13	W	0.412*	0.430	0.456	0.477	0.533	0.550	0.613	0.696	0.775
1/4	1	-	P	0.324	0.337	0.378	0.423	0.470	0.533	0.648	0.791	0.965
		0.05	UW	0.313-	0.325	0.349	0.383	0.434	0.459	0.553	0.663	0.859
		0.41	W	0.329	0.349	0.381*	0.420	0.469	0.509	0.616	0.745	0.928
1/4	3	-	P	0.208	0.229	0.239	0.245	0.278	0.293	0.352	0.402	0.514
		0.38	UW	0.216	0.213	0.232	0.247*	0.269	0.289	0.342	0.389	0.500
		1.00	W	0.230	0.231	0.253	0.270	0.297	0.319	0.376	0.439	0.559+
1/2	1/3	-	P	0.418	0.441	0.469	0.510	0.549	0.624	0.683	0.795	0.898
		0.19	UW	0.429	0.451*	0.459	0.484	0.496	0.549	0.576	0.645	0.721
		0.24	W	0.431	0.459*	0.468	0.491	0.512	0.566	0.596	0.671	0.751
1/2	1	-	P	0.310	0.360	0.375	0.418	0.474	0.538	0.642	0.790	0.969
		0.43	UW	0.343	0.369	0.394*	0.416	0.456	0.517	0.594	0.696	0.885
		0.53	W	0.344	0.379	0.403	0.428*	0.474	0.533	0.619	0.733	0.913
1/2	3	-	P	0.204	0.205	0.235	0.246	0.281	0.296	0.347	0.410	0.510
		1.00	UW	0.239	0.244	0.270	0.290	0.321	0.333	0.387	0.464	0.564+
		1.00	W	0.240	0.247	0.275	0.297	0.328	0.348	0.401	0.487	0.594+
1	1/3	-	P	0.406	0.453	0.473	0.512	0.558	0.621	0.707	0.780	0.894
		0.31	UW	0.438	0.472	0.484*	0.502	0.527	0.564	0.605	0.645	0.719
		0.31	W	0.438	0.472	0.484*	0.502	0.527	0.564	0.605	0.645	0.719
1	1	-	P	0.313	0.338	0.373	0.414	0.468	0.534	0.638	0.793	0.971
		0.62	UW	0.379	0.401	0.420	0.458	0.490	0.549*	0.626	0.738	0.909
		0.62	W	0.379	0.401	0.420	0.458	0.490	0.549*	0.626	0.738	0.909
1	3	-	P	0.197	0.211	0.230	0.245	0.264	0.307	0.339	0.406	0.507
		1.00	UW	0.270	0.290	0.306	0.320	0.348	0.398	0.443	0.527	0.671+
		1.00	W	0.270	0.290	0.306	0.320	0.348	0.398	0.443	0.527	0.671+
2	1/3	-	P	0.412	0.456	0.463	0.523	0.563	0.616	0.694	0.784	0.889
		0.38	UW	0.456	0.476	0.477*	0.515	0.537	0.562	0.610	0.660	0.717
		0.36	W	0.455	0.475	0.473*	0.510	0.530	0.552	0.599	0.644	0.699
2	1	-	P	0.317	0.341	0.368	0.413	0.472	0.537	0.642	0.810	0.964
		0.71	UW	0.399	0.428	0.445	0.483	0.529	0.575	0.652*	0.769	0.918
		0.68	W	0.401	0.428	0.446	0.479	0.521	0.567*	0.638	0.749	0.902
2	3	-	P	0.207	0.218	0.236	0.260	0.275	0.300	0.341	0.395	0.508
		1.00	UW	0.316	0.322	0.346	0.373	0.398	0.450	0.518	0.616	0.776+
		1.00	W	0.314	0.320	0.347	0.371	0.391	0.447	0.512	0.597	0.759+

^a ρ_T : theoretical maximum ρ 's for the unweighted or weighted P-value pooling test T_1 being more powerful than the naive paired test T_p when data is normally distributed and variances are known.

^b **P**: naive paired test; **UW**: unweighted combination test; **W**: weighted combination test.

^c *: the estimated maximum ρ 's for the unweighted or weighted P-value pooling test T_1 being more powerful than the naive paired test T_p is between corresponding ρ and $\rho + 0.1$; -: estimated maximum ρ is less than 0.1; +: estimated maximum ρ is larger than 0.9.

Table S2: With incompleteness in both arms under bivariate normal distribution with unknown variances:
power of naive paired

test T_p , unweighted and weighted combination tests T_2 given $\gamma_1 = n_2/n_1, \gamma_2 = n_3/n_2 = 1, \nu^2 = \sigma_X^2/\sigma_Y^2$, and ρ .

γ_1	ν^2	ρ_T	Method	ρ								
				0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1/4	1/3	-	P	0.413	0.439	0.466	0.508	0.554	0.627	0.696	0.777	0.894
		0.36	UW	0.424	0.461	0.482	0.510*	0.532	0.583	0.634	0.692	0.798
		0.90	W	0.466	0.498	0.530	0.567	0.605	0.664	0.723	0.788*	0.889
1/4	1	-	P	0.313	0.350	0.374	0.415	0.471	0.525	0.658	0.803	0.971
		0.31	UW	0.334	0.345	0.376*	0.405	0.432	0.493	0.575	0.686	0.877
		0.78	W	0.358	0.385	0.417	0.453	0.496	0.563	0.670*	0.795	0.959
1/4	3	-	P	0.211	0.215	0.236	0.240	0.270	0.314	0.350	0.413	0.511
		0.36	UW	0.229	0.222	0.248	0.244*	0.266	0.278	0.315	0.364	0.434
		0.90	W	0.243	0.236	0.267	0.265	0.295	0.322	0.367	0.426	0.516+
1/2	1/3	-	P	0.411	0.430	0.463	0.510	0.574	0.623	0.698	0.784	0.883
		0.76	UW	0.512	0.537	0.559	0.594	0.614	0.650	0.710*	0.773	0.856
		0.92	W	0.528	0.550	0.578	0.618	0.648	0.689	0.748	0.821	0.894+
1/2	1	-	P	0.321	0.342	0.375	0.426	0.463	0.547	0.640	0.804	0.969
		0.66	UW	0.406	0.416	0.463	0.480	0.504	0.562*	0.634	0.754	0.906
		0.80	W	0.417	0.429	0.479	0.502	0.527	0.602	0.681	0.807*	0.944
1/2	3	-	P	0.220	0.212	0.229	0.244	0.270	0.296	0.346	0.406	0.511
		0.76	UW	0.255	0.255	0.271	0.286	0.306	0.323	0.354*	0.403	0.477
		0.92	W	0.266	0.261	0.283	0.295	0.322	0.338	0.374	0.435	0.514+
1	1/3	-	P	0.404	0.453	0.463	0.514	0.552	0.621	0.694	0.795	0.886
		0.96	UW	0.636	0.657	0.672	0.694	0.724	0.760	0.803	0.855	0.915+
		0.96	W	0.636	0.657	0.672	0.694	0.724	0.760	0.803	0.855	0.915+
1	1	-	P	0.309	0.349	0.356	0.416	0.468	0.544	0.640	0.789	0.969
		0.83	UW	0.486	0.511	0.533	0.555	0.599	0.654	0.722	0.822*	0.938
		0.83	W	0.486	0.511	0.533	0.555	0.599	0.654	0.722	0.822*	0.938
1	3	-	P	0.210	0.215	0.233	0.250	0.275	0.297	0.346	0.410	0.501
		0.96	UW	0.318	0.320	0.332	0.348	0.359	0.397	0.428	0.470	0.537+
		0.96	W	0.318	0.320	0.332	0.348	0.359	0.397	0.428	0.470	0.537+
2	1/3	-	P	0.399	0.429	0.482	0.526	0.551	0.618	0.689	0.785	0.892
		1.00	UW	0.773	0.792	0.823	0.826	0.843	0.860	0.892	0.925	0.964+
		1.00	W	0.785	0.794	0.824	0.827	0.840	0.849	0.879	0.905	0.950+
2	1	-	P	0.330	0.345	0.380	0.414	0.464	0.549	0.648	0.804	0.966
		0.92	UW	0.636	0.659	0.658	0.687	0.714	0.766	0.826	0.894	0.971+
		0.87	W	0.645	0.660	0.653	0.675	0.702	0.744	0.800	0.863*	0.947
2	3	-	P	0.211	0.211	0.218	0.238	0.271	0.303	0.347	0.398	0.505
		1.00	UW	0.399	0.401	0.413	0.435	0.470	0.487	0.527	0.557	0.621+
		1.00	W	0.403	0.403	0.421	0.439	0.464	0.478	0.510	0.530	0.578+

^a ρ_T : theoretical maximum ρ 's for the unweighted or weighted P-value pooling test T_1 being more powerful than the naive paired test T_p when data is normally distributed and variances are known.

^b **P**: naive paired test; **UW**: unweighted combination test; **W**: weighted combination test.

^c *: the estimated maximum ρ 's for the unweighted or weighted P-value pooling test T_1 being more powerful than the naive paired test T_p is between corresponding ρ and $\rho + 0.1$; -: estimated maximum ρ is less than 0.1; +: estimated maximum ρ is larger than 0.9.

Table S3: With incompleteness in single arm under bivariate logistic distribution: power of naive paired test T_p , unweighted and weighted combination tests T_1 given $\gamma_1 = n_2/n_1$, $\nu^2 = \sigma_X^2/\sigma_Y^2$, and ρ .

γ_1	ν^2	ρ_T	Method	ρ								
				0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1/4	1/3	-	P	0.402	0.410	0.433	0.461	0.489	0.511	0.549	0.586	0.632
		0	UW	0.392-	0.394	0.409	0.428	0.438	0.450	0.477	0.495	0.517
		0.13	W	0.401-	0.407	0.423	0.443	0.452	0.474	0.500	0.522	0.550
1/4	1	-	P	0.284	0.302	0.330	0.363	0.408	0.462	0.550	0.710	0.924
		0.05	UW	0.296*	0.301	0.320	0.334	0.381	0.421	0.483	0.605	0.807
		0.41	W	0.308	0.309	0.342	0.361	0.412*	0.455	0.530	0.669	0.878
1/4	3	-	P	0.132	0.134	0.127	0.134	0.141	0.158	0.151	0.156	0.174
		0.38	UW	0.145	0.137	0.147	0.152	0.160	0.166	0.175	0.179	0.196+
		1.00	W	0.147	0.144	0.150	0.158	0.162	0.169	0.180	0.186	0.205+
1/2	1/3	-	P	0.406	0.409	0.438	0.465	0.491	0.527	0.540	0.580	0.631
		0.19	UW	0.397-	0.404	0.416	0.433	0.456	0.473	0.475	0.490	0.515
		0.24	W	0.399-	0.405	0.421	0.438	0.462	0.479	0.486	0.502	0.532
1/2	1	-	P	0.282	0.305	0.332	0.364	0.414	0.467	0.567	0.708	0.925
		0.43	UW	0.307	0.319	0.343	0.369	0.419*	0.449	0.520	0.633	0.831
		0.53	W	0.313	0.327	0.351	0.377	0.430	0.471*	0.544	0.660	0.859
1/2	3	-	P	0.126	0.132	0.130	0.137	0.137	0.136	0.153	0.157	0.175
		1.00	UW	0.151	0.141	0.148	0.156	0.165	0.172	0.171	0.179	0.201+
		1.00	W	0.154	0.146	0.151	0.159	0.166	0.174	0.177	0.179	0.206+
1	1/3	-	P	0.404	0.421	0.448	0.460	0.490	0.512	0.548	0.580	0.636
		0.31	UW	0.411*	0.419	0.442	0.437	0.470	0.462	0.477	0.505	0.526
		0.31	W	0.411*	0.419	0.442	0.437	0.470	0.462	0.477	0.505	0.526
1	1	-	P	0.287	0.309	0.339	0.364	0.404	0.477	0.575	0.712	0.928
		0.62	UW	0.339	0.356	0.378	0.408	0.434	0.491*	0.560	0.659	0.857
		0.62	W	0.339	0.356	0.378	0.408	0.434	0.491*	0.560	0.659	0.857
1	3	-	P	0.128	0.133	0.136	0.127	0.132	0.140	0.167	0.160	0.164
		1.00	UW	0.161	0.171	0.160	0.169	0.175	0.195	0.209	0.212	0.225+
		1.00	W	0.161	0.171	0.160	0.169	0.175	0.195	0.209	0.212	0.225+
2	1/3	-	P	0.397	0.404	0.445	0.463	0.492	0.511	0.548	0.593	0.644
		0.38	UW	0.405	0.410*	0.426	0.447	0.464	0.464	0.482	0.510	0.540
		0.36	W	0.404	0.409*	0.425	0.442	0.462	0.460	0.475	0.503	0.531
2	1	-	P	0.286	0.302	0.328	0.369	0.417	0.458	0.560	0.702	0.922
		0.71	UW	0.359	0.375	0.381	0.421	0.450	0.498	0.567*	0.682	0.860
		0.68	W	0.360	0.373	0.381	0.423	0.447	0.491*	0.557	0.670	0.836
2	3	-	P	0.132	0.125	0.122	0.138	0.137	0.150	0.159	0.165	0.164
		1.00	UW	0.189	0.181	0.187	0.208	0.216	0.224	0.254	0.259	0.303+
		1.00	W	0.194	0.180	0.191	0.206	0.219	0.224	0.253	0.257	0.298+

^a ρ_T : theoretical maximum ρ 's for the unweighted or weighted P-value pooling test T_1 being more powerful than the naive paired test T_p when data is normally distributed and variances are known.

^b **P**: naive paired test; **UW**: unweighted combination test; **W**: weighted combination test.

^c *: the estimated maximum ρ 's for the unweighted or weighted P-value pooling test T_1 being more powerful than the naive paired test T_p is between corresponding ρ and $\rho + 0.1$; -: estimated maximum ρ is less than 0.1; +: estimated maximum ρ is larger than 0.9.

Table S4: With incompleteness in both arms under bivariate logistic distribution: power of naive paired test T_p , unweighted and weighted combination tests T_2 given $\gamma_1 = n_2/n_1, \gamma_2 = n_3/n_2 = 1, \nu^2 = \sigma_X^2/\sigma_Y^2$, and ρ .

γ_1	ν^2	ρ_T	Method	ρ									
				0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
1/4	1/3	-	P	0.404	0.431	0.441	0.455	0.474	0.512	0.551	0.585	0.634	
				0.36	UW	0.445	0.463	0.459	0.473	0.488	0.522*	0.550	0.564
				0.90	W	0.473	0.495	0.498	0.517	0.533	0.573	0.600	0.626
1/4	1	-	P	0.271	0.301	0.340	0.357	0.404	0.464	0.577	0.699	0.928	
				0.31	UW	0.285	0.325	0.349	0.361*	0.393	0.433	0.514	0.610
				0.78	W	0.310	0.345	0.376	0.395	0.439	0.493	0.601	0.704*
1/4	3	-	P	0.117	0.134	0.129	0.134	0.141	0.145	0.152	0.163	0.164	
				0.36	UW	0.130	0.144	0.145	0.146	0.150	0.148	0.155	0.168*
				0.90	W	0.133	0.151	0.147	0.152	0.157	0.157	0.162	0.177
1/2	1/3	-	P	0.398	0.419	0.439	0.463	0.478	0.526	0.544	0.578	0.626	
				0.76	UW	0.510	0.532	0.537	0.556	0.561	0.598	0.610	0.629
				0.92	W	0.520	0.546	0.548	0.567	0.581	0.622	0.634	0.654
1/2	1	-	P	0.290	0.303	0.332	0.353	0.405	0.462	0.565	0.709	0.925	
				0.66	UW	0.357	0.370	0.379	0.400	0.446	0.485	0.567*	0.662
				0.80	W	0.362	0.380	0.397	0.417	0.462	0.515	0.601	0.716*
1/2	3	-	P	0.131	0.123	0.136	0.126	0.137	0.145	0.160	0.157	0.163	
				0.76	UW	0.156	0.141	0.155	0.150	0.154	0.165	0.177	0.168
				0.92	W	0.155	0.142	0.158	0.147	0.158	0.172	0.185	0.173
1	1/3	-	P	0.401	0.411	0.443	0.461	0.478	0.506	0.545	0.582	0.621	
				0.96	UW	0.612	0.623	0.635	0.652	0.669	0.676	0.706	0.726
				0.96	W	0.612	0.623	0.635	0.652	0.669	0.676	0.706	0.726
1	1	-	P	0.281	0.305	0.329	0.356	0.403	0.479	0.564	0.704	0.929	
				0.83	UW	0.428	0.446	0.467	0.494	0.515	0.575	0.645	0.728*
				0.83	W	0.428	0.446	0.467	0.494	0.515	0.575	0.645	0.728*
1	3	-	P	0.127	0.135	0.131	0.133	0.145	0.148	0.153	0.156	0.167	
				0.96	UW	0.172	0.179	0.169	0.180	0.181	0.185	0.193	0.195
				0.96	W	0.172	0.179	0.169	0.180	0.181	0.185	0.193	0.195
2	1/3	-	P	0.396	0.423	0.429	0.459	0.490	0.516	0.547	0.585	0.641	
				1.00	UW	0.748	0.752	0.766	0.782	0.790	0.806	0.821	0.848
				1.00	W	0.756	0.754	0.772	0.788	0.787	0.807	0.819	0.840
2	1	-	P	0.280	0.305	0.328	0.373	0.410	0.467	0.571	0.709	0.923	
				0.92	UW	0.528	0.557	0.572	0.605	0.641	0.675	0.738	0.811
				0.87	W	0.534	0.559	0.579	0.597	0.633	0.659	0.713	0.780*
2	3	-	P	0.118	0.134	0.133	0.132	0.146	0.141	0.151	0.153	0.166	
				1.00	UW	0.194	0.205	0.215	0.208	0.216	0.224	0.231	0.227
				1.00	W	0.198	0.205	0.205	0.206	0.212	0.218	0.226	0.228

^a ρ_T : theoretical maximum ρ 's for the unweighted or weighted P-value pooling test T_1 being more powerful than the naive paired test T_p when data is normally distributed and variances are known.

^b **P**: naive paired test; **UW**: unweighted combination test; **W**: weighted combination test.

^c *: the estimated maximum ρ 's for the unweighted or weighted P-value pooling test T_1 being more powerful than the naive paired test T_p is between corresponding ρ and $\rho + 0.1$; -: estimated maximum ρ is less than 0.1; +: estimated maximum ρ is larger than 0.9.